4

Performance Tests of Physical Ability

DORLY J.H. DEEG[1]

Physical ability is a major indicator of the health status of older adults. Rather than specifying pathological conditions, it indicates the severity, stage, and possible sequelae to pathology. This has long been considered indispensable in clinical evaluation (Katz *et al.* 1963). Moreover, from the perspective of older persons themselves, quality of life is more closely related to physical ability than to underlying pathology. Also, physical ability gives more information on care needs than the specification of pathologic conditions (Guralnik *et al.* 1989).

Measurement instruments that are most often used to assess physical ability in surveys of the general population consist of questionnaires on basic and instrumental activities of daily living and on functional limitations. However, the self-report character of these instruments may imply that instead of actual abilities, the respondents' perception of their abilities is studied. This problem has been recognized by the scientific community. As a consequence, performance tests of physical ability are increasingly included in home surveys of older persons to complement self-reports of physical ability (Magaziner and Guralnik 1992). Performance tests determine the actual, momentaneous ability to perform tasks. These tests are internationally in an experimental phase. Research is in progress to assess when performance measures offer advantages over self-report measures, particularly with regard to sensitivity to change. Two recent studies of performance measures vs. self-reports showed that performance test scores predict the outcomes of death or nursing home placement - independent of self-reports (Reuben *et al.* 1992, Guralnik *et al.* 1994). A third study demonstrated that three-year change in physical performance was predicted by health parameters such as lung function (peak flow) and admission to hospital in a high-functioning older cohort (Seeman *et al.* 1994).

---

[1] Dorly J.H. Deeg, Ph.D., is Associate Professor in the Department of Psychiatry, Vrije Universiteit, and Scientific Director of the Longitudinal Aging Study Amsterdam. Her research interest is the epidemiology of aging.

A selection of performance tests was included in LASA to indicate both upper and lower body functioning. This chapter evaluates their feasibility in a home survey of older persons in terms of non-participation rate and non-participant characteristics. Factors affecting participation are expected in the realms of both physical and emotional functioning. This chapter continues to describe the distribution of performance across age and sex, and the association of performance tests with self-reports of physical ability.

Methods

Sample

Both the physical performance tests and the self-reports of physical ability were included in the main interview. The self-reports and the tests of upper body functioning were also included in the short version of the interview; the tests for lower body functioning were administered in the complete interview only (n = 2936).

Table 1
Four tests of physical performance and underlying parameters of functional ability

	Underlying parameter			
	Mobility	Balance	Coordination	Strength
UPPER BODY				
Put on/take off cardigan	X		X	
Close button	X		X	
LOWER BODY				
Walk 6 m	X	X		
Chair stands		X		X

Procedures

Performance tests were selected as measures of mobility, balance, coordination, and strength (see table 1). Mobility is taken to include upper-body flexibility, and finger dexterity. Performance was measured by the following tasks: (1) to put on and take off a cardigan which was brought in by the interviewer, (2) to close one specific button when the cardigan was on, (3) to walk three meters back and forth along a rope,

and (4) to get up from a kitchen chair five times arms folded (NCHS 1990, Magaziner 1991, Reuben and Siu 1990). The time needed to do each task was measured in seconds using a stopwatch. Time needed has been shown to correlate well with physical ability (Spiegel *et al.* 1987). During the walk, the interviewer recorded the number of steps taken. Irregularities in the performance of each test were recorded as dummy variables (Tinetti *et al.* 1986).

The interviewers were trained for this part of the interview using an instruction video tape produced by the U.S. National Institute on Aging. Because these tests involve some risk for the participants, the interviewers were instructed not to do a test if they judged it not safe or if a participant objected to it.

In order to avoid analyzing a positive selection of the sample, each subject having a time score on a performance test was gives a score 1 through 4 corresponding to the quartile of the distribution of time needed: the more time was needed, the higher the score. Those subjects who did not complete the test, but did answer the questions on functional limitations, were given the score 5 (cf. Guralnik *et al.* 1994). Those who did not answer all questions on functional limitations remained excluded from analyses (n = 20). The internal consistency of the four new items was good, given the small number of items (Cronbach's alpha = 0.62). The four items were summed to a performance summary score.

Self-reports on physical ability were obtained on the following items: using transportation, going up and down a staircase, and cutting one's toenails. The respondents were asked to indicate whether they had difficulty doing the activity, whether they needed help, or were not able to do it at all. An index of functional limitations was created by summing the items for which the respondent reported difficulty or need of help.

Emotional functioning was measured during the main interview using the Center for Epidemiologic Studies Depression scale (Beekman *et al.* 1994).

Statistical methods

Participants and non-participants were compared using chi-square tests for categorical variables and t-tests for continuous variables. Distributions of time scores and number of steps across age and sex were tested using analysis of variance. The association of performance summary score with self-reports of physical ability was tested sex-specifically using linear regression models of performance on self-report score including age as a control variable.

Results

Among respondents, the participation rate was 97% for the cardigan, 96% for the button, 97% for the walk, and 91% for the chair stands, respectively. The chair stands was the test most often considered not safe or too taxing for the respondent. Non-participants in all tests were older, reported poorer functional ability, and more often reported depressive symptoms (table 2). Non-participants in tests of upper body functioning were more often male. There were no sex differences in participation in tests of lower body functioning.

Table 2
Characteristics of participants (P) vs. non-participants (NP) on each performance test

	% ≥ 70 jr		% Male		% ≥ 1 FL[1]		Mean CES-D[2]	
	P	NP	P	NP	P	NP	P	NP
UPPER BODY								
Cardigan on	52.2	72.5	48.2	58.2	40.4	82.3	7.8	15.9
Cardigan off	52.2	73.2	48.2	57.7	40.4	80.0	7.8	15.4
Close button	52.0	73.0	48.0	60.3	39.9	81.6	7.8	14.4
LOWER BODY								
Walk 6 m time	50.6	76.0	48.7	46.1	38.7	72.3	7.6	12.9
Walk 6 m steps	50.6	76.5	48.6	46.5	38.2	72.0	7.6	12.9
Chair stands	49.4	74.5	49.0	44.9	35.7	79.5	7.3	13.0

[1] FL, functional limitations
[2] CES-D, Center for Epidemiologic Studies Depression scale
Note: All comparisons between P and NP are significant ($p < 0.05$) except the percent males in the lower body tests.

Table 3 shows that mean and median of the tests may have quite different values. The distribution of all tests is skewed to the right. Since the correlation of putting on and taking off the cardigan is reasonably high, the two time scores are summed for further analyses.

Table 3
Descriptive characteristics of performance tests
The time is indicated in seconds

	Mean	Median	St. dev.	Percentile 25	Percentile 75
UPPER BODY					
Put on cardigan time[1]	8.7	7.0	6.2	5.0	10.0
Take off cardigan time[1]	5.3	5.0	3.8	4.0	6.0
Cardigan on + off time	14.0	12.0	8.5	9.0	16.0
Close button time	6.6	4.0	8.3	2.0	8.0
LOWER BODY					
Walk 6 m time	8.6	7.0	5.9	6.0	9.0
Walk 6 m steps	11.7	11.0	4.4	9.0	13.0
Chair stands time	12.7	12.0	5.0	10.0	14.0

[1] correlation (on, off) = 0.40.

Interviewer observations of irregularities in each of the tests are listed in table 4. In the cardigan test, help was needed most often with taking off the cardigan. In the button test, irregularities occurred mostly in choosing the button or the buttonhole (4%). For the walk, over 3% of respondents used an aid. Also, quite many respondents were observed to have gait irregularities. The most prevalent irregularity was instability (9%). Similarly, relatively many respondents were unstable while standing up doing the chair stands (6%). Since all irregularities were associated with a greater time needed to complete the test, irregularities are considered to be reflected in the time score, and disregarded in further analyses.

Age and sex differences in performance on each test are shown in table 5. In both males and females, time needed increased across age groups. For most tests, the increase was non-linear, and accelerated with increasing age (all age associations were significant at $p < 0.001$). The sexes performed differently on each test. In all age groups, women needed less time on the cardigan and button tests, and men needed less time on the lower body tests. The sex differences in performance were progressively greater in older age groups (the interaction term of age and sex in analysis of variance was significant at $p < 0.05$). The performance summary score also showed age and sex differences. However, sex differences were of similar size in all age groups (the interaction term was not significant).

Table 4
Observation of irregularities in performance tests in those who completed each test

	% irregular
CARDIGAN ON	
First arm with help	0.8
Across shoulder with help	1.5
Second arm with help	1.4
CARDIGAN OFF	
First arm free with help	3.6
Second arm free with help	2.6
BUTTON	
Grasping with help	0.7
Through hole with help	0.5
Wrong button or wrong buttonhole	4.5
WALK	
Walking aids	3.2
Pain during walk	8.2
Gait: problems starting	4.1
unstable gait	8.8
limp	0.2
stiff legs	5.3
leg slides	3.6
training leg	1.2
unstable turn	7.0
irregular gait	7.9
other	2.4
CHAIR STANDS	
Wheelchair	0.3
Sliding forward	2.5
Use of arms	2.6
Number of times less than five	1.5
Unstable during rise	6.0

The association of performance with self-reported physical ability was considerable, and greater than the association with age (table 6, $p < 0.001$). In men, the self-reports added 13% to the variance explained by age in the performance summary score; in women, self-reports added even 17%. The association between self-reports and performance seemed greater in women than in men. However, this difference was not significant (the interaction term of sex and functional limitations in a regression model including both sexes did not reach significance). Interestingly, the strength of the association of self-reports with performance was reduced in older age groups (the interaction term of age and functional limitations was significant for both men, $p = 0.05$, and women, $p < 0.01$).

Table 5
Associations of single performance tests and summary score with age and sex

	Age in years					
	55-59	60-64	65-69	70-74	75-79	80-85
Males						
UPPER BODY						
Cardigan on + off time	12.3	13.3	13.7	15.2	16.9	20.8
Button time	5.1	6.9	7.3	8.7	9.5	16.3
LOWER BODY						
Walk time	6.5	7.1	7.4	7.7	8.5	11.1
Walk steps	9.4	9.4	9.9	10.4	11.1	13.5
Chair stands time	10.4	11.2	11.9	12.3	14.0	15.5
Summary score	8.6	9.3	9.6	10.6	12.0	13.4
Females						
UPPER BODY						
Cardigan on + off time	9.8	10.5	11.4	12.4	13.8	16.6
Button time	3.1	3.4	3.6	4.0	4.9	6.2
LOWER BODY						
Walk time	6.7	7.5	7.9	8.8	10.8	12.5
Walk steps	10.4	11.1	11.9	12.3	14.0	15.5
Chair stands time	11.1	11.7	12.7	13.2	15.1	15.8
Summary score	7.4	7.9	9.2	10.1	11.3	12.8

Table 6
Association of performance summary score with functional limitations self-reports by sex. Linear regression with dependent variable: performance summary score

	Standardized regression coefficient	
	Males	Females
Age	0.33	0.25
Functional limitations	0.38	0.49
Variance explained by age	21.1%	25.3%
Total variance explained	34.1%	42.7%

Discussion

From the findings described above, the administration of performance tests in home surveys of older persons proves to be feasible. The participation rate was satisfactory, and even better than rates reported elsewhere (Guralnik *et al.* 1994). The findings regarding non-participation in the tests suggest that the decision to stay on the safe side administering performance tests is based on both physical and mental conditions. Although the interviewers were carefully trained to maintain standard procedures, the amount of encouragement given, especially in depressed respondents, may affect both the participation rate and the test outcome (Guyatt *et al.* 1984). This is an area for further research.

As compared to women, among men the participation rate in the cardigan and button tests was notably lower, and their performance on these tests poorer. By contrast, ample evidence shows that women in all age groups are less physically able than men. Our findings may be partly explained by the unfamiliarity of men with an article of clothing such as a cardigan. To make things worse, the button was on the 'female' side (right over left) on all cardigans used. Analyses involving these upper body tests should be carried out for each sex separately.

The association of self-reports and performance tests of physical ability was considerable, given the differences in content and mode of measurement. However, the variance explained was less than 20%, indicating that only part of the underlying physical ability is measured by both instruments. Therefore, the performance tests selected in LASA seem to be a valuable addition to self-reports on physical ability. Additional research could address the issue of which factors are related to discrepancies in self-reports and performance, especially in the oldest-old. Furthermore, the

selection of tests should be validated against more extensive measurements of physical performance in order to evaluate what range of the total spectrum of physical functioning is covered.

References

Beekman ATF, Limburg J van, Deeg DJH, Tilburg W van, Wouters L (1994) [Screening for depression in the elderly in the community: Using the Center for Epidemiologic Studies Depression Scale (CES-D) in the Netherlands], *Tijdschr Gerontol Geriatr* 25: 95-103 (in Dutch, summary in English).

Guralnik JM, Branch LG, Cummings SR, Curb JD (1989) Physical performance measures in aging research, *J Gerontol* 44: M141-146.

Guralnik JM, Simonsick EM, Ferrucci L, Glynn RJ, Berkman LF, Blazer DG, Scherr PA, Wallace RB (1994) A short physical performance battery assessing lower extremity function: Association with self-reported disability and prediction of mortality and nursing home admission, *J Gerontol* 49: M85-94.

Guyatt GH, Pugsley SO, Sullivan MJ, Thompson PJ, Berman LB, Jones NL, Fallen EL, Taylor DW (1984) Effect of encouragement on walking test performance. *Thorax* 39: 818-822.

Katz SC, Ford AB, Moskowitz RW, Jackson BA, Jaffe MW (1963) Studies of illness in the aged. The index of ADL: A standardized measure of biological and psychosocial function, *J Am Med Assoc* 185: 914-919.

Magaziner J (1991) *Hip Fracture Recovery Study Patient-Proxy Concordance Study Field Manual*, Baltimore, MD, USA: University of Maryland School of Medicine.

Magaziner J, Guralnik J (1992) The utility of performance measures of physical functioning: insights from ongoing studies, *Gerontologist* 32 [Spec Issue II]: 88.

National Center for Health Statistics (NCHS) (1990) *National Health and Nutrition Examination Survey III data collection forms*, Hyattsville, MD, USA: US Department of Health and Human Services.

Reuben DB, Siu AL (1990) An objective measure of physical function of elderly outpatients: The Physical Performance Test, *J Am Geriatr Soc* 38: 1105-1112.

Reuben DB, Siu AL, Kimpau S (1992) The predictive validity of self-report and performance-based measures of function and health, *J Gerontol* 47: M106-110.

Seeman TE, Charpentier PA, Berkman LF, Tinetti ME, Guralnik JM, Albert M, Blazer D, Rowe JW (1994) Predicting changes in physical performance in a high-functioning elderly cohort: MacArthur studies on successful aging, *J Gerontol* 49: M97-108.

Spiegel JS, Paulus HE, Ward NB, Spiegel TM, Leake B, Kane RL (1987) What are we measuring? An examination of walk time and grip strength, *J Rheumatol* 14: 80-86.

Tinetti ME, Williams TF, Mayewski R (1986) Fall risk index for elderly patients based on number of chronic disabilities, *Am J Med* 80: 429-434.